

# **Outsourcing TA-4J Maintenance: Cost and Quality Experience**

LCDR Christopher Reeger

**DISTRIBUTION STATEMENT A**  
Approved for Public Release  
Distribution Unlimited

**Center for Naval Analyses**

4401 Ford Avenue • Alexandria, Virginia 22302-1498

**DTIC QUALITY INSPECTED 2**

19990429 013

Approved for distribution:

J



Samuel D. Kleinman, Director  
Infrastructure and Readiness Team  
Support Planning and Management Division

This document represents the best opinion of CNA at the time of issue.  
It does not necessarily represent the opinion of the Department of the Navy.

**APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED**

For copies of this document, call the CNA Document Control and Distribution Section (703) 8

**REPORT DOCUMENTATION PAGE**Form Approved  
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

<b>1. AGENCY USE ONLY (Leave blank)</b>		<b>2. REPORT DATE</b> July 1997	<b>3. REPORT TYPE AND DATES COVERED</b> Final	
<b>4. TITLE AND SUBTITLE</b> Outsourcing TA-4J Maintenance: Cost and Quality Experience			<b>5. FUNDING NUMBERS</b> C - N00014-96-D-0001	
<b>6. AUTHOR(S)</b> CM Reeger				
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Center for Naval Analyses 4401 Ford Avenue Alexandria, Virginia 22302-1498			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b> CRM 97-30	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>			<b>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</b>	
<b>11. SUPPLEMENTARY NOTES</b>				
<b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b> Distribution unlimited			<b>12b. DISTRIBUTION CODE</b>	
<b>13. ABSTRACT (Maximum 200 Words)</b> Despite shrinking budgets, the U.S. military is struggling to simultaneously fund force levels, current operations, and an aggressive modernization program. Many believe the military can fund its recapitalization program if cost efficiencies can be achieved from within infrastructure budgets. One way to reduce infrastructure costs is through competition, outsourcing, and privatization. Whether the in-house (or organic) team or the private team wins the contract, the government benefits because the competition lowers costs and increases productivity. This paper examines the maintenance of the Navy's TA-4Js. The value of this analysis is that it allows us to look at a long series of performance and cost data, both for in-house and contractor maintenance. Because we have data on three contractors, we can also examine the effect of changing contractors.				
<b>14. SUBJECT TERMS</b> aircraft maintenance, competition, costs, defense economics, fully mission capable (FMC), maintenance, methodology, mission capable rate (MCR), outsourcing, quality, Skyhawk TA-4J (naval training aircraft), training			<b>15. NUMBER OF PAGES</b> 45	
			<b>16. PRICE CODE</b>	
<b>17. SECURITY CLASSIFICATION OF REPORT</b> Unclassified	<b>18. SECURITY CLASSIFICATION OF THIS PAGE</b> Unclassified	<b>19. SECURITY CLASSIFICATION OF ABSTRACT</b> Unclassified	<b>20. LIMITATION OF ABSTRACT</b>	

# Contents

<b>Summary</b> . . . . .	1
<b>Introduction</b> . . . . .	5
<b>Background</b> . . . . .	7
<b>Analysis</b> . . . . .	11
Intermediate-level maintenance. . . . .	18
Break-in period. . . . .	21
<b>Productivity</b> . . . . .	25
Labor productivity versus technological advances . . . . .	27
<b>Conclusions</b> . . . . .	29
<b>Appendix: Methodology</b> . . . . .	31
<b>References</b> . . . . .	35
<b>List of figures</b> . . . . .	37
<b>List of tables</b> . . . . .	39

## Summary

Despite shrinking budgets, the U.S. military is struggling to simultaneously fund force levels, current operations, and an aggressive modernization program. Many believe the military can fund its recapitalization program if cost efficiencies can be achieved from within infrastructure budgets. One way to reduce infrastructure costs is through competition, outsourcing, and privatization. Whether the in-house (or organic) team or the private team wins the contract, the government benefits because the competition lowers costs and increases productivity.

Outsourcing, or more precisely, the competition and cost visibility brought about by outsourcing, has been a successful weapon in the fight to reduce costs [1]. Warfighters want to know: "Will we have the support to complete our mission?" The current infrastructure gets the job done. Will squeezing the infrastructure for savings put the Navy at too high a risk? To meet its mission, the fleet must have proper training, qualified personnel, superior equipment, and other infrastructure products. These products need not be provided wholly by government employees.

This paper examines the maintenance of the Navy's TA-4Js. The value of this analysis is that it allows us to look at a long series of performance and cost data, both for in-house and contractor maintenance. Because we have data on three contractors, we can also examine the effect of changing contractors.

The Navy has flown A-4 Skyhawk light attack aircraft since the 1950s. For more than 25 years, the Naval Aviation Training Command has used a two-seat version of the jet, the TA-4J, to train student naval aviators in the advanced phase of strike training. Until the mid 1980s, the Navy used its own personnel and equipment for the organizational-(O-) and intermediate-(I-) level maintenance of the aircraft. At

that point, the decision was made to open the maintenance to private companies.

Lockheed won the first competition in 1985 for O-level maintenance with a bid that was about 20 percent lower than the in-house offer. In 1989, Burnside OTT won the contract for I-level maintenance. In 1990, Grumman replaced Lockheed as the contractor for O-level. Then in 1993, O-level and I-level maintenance were combined into one contract for both the TA-J and the T-2 and awarded to UNC. This fixed-price contract consists of 1 base year and 4 option years. Finally, in 1995 Sabreliner was awarded the contract for depot- (D-) level maintenance of both platforms. The TA-4J portion of this contract was estimated at about \$11.3 million.

We analyzed a range of data, but we focused on direct maintenance man-hours per flight hour (DMMH), full mission capable rate (FMC), and mission capable rate (MC). DMMH is one measure of the labor input, and MC and FMC are both measures of performance.

Two themes emerged from the analysis. First, contractors performed better than the Navy in-house team in almost every case. For both maintenance man-hour and FMC rates, the contractor means were considered "better." The means of the MC rates for both the Navy in-house team and the contractors were about equal. This seems to show that the contractors were able to achieve a similar MC rate and a better FMC rate than those of the in-house Navy maintenance team—and do it in fewer man-hours. Here, outsourcing results in the same or better quality product and increased efficiency.

The second finding shows that at the start of the contract, performance, as measured by the maintenance statistics, declined. In almost all cases, the contractor's initial performance was worse than that of the in-house team, or it exceeded the in-house Navy standard only to deteriorate thereafter. On average, it was about 2 years before the contractor began to improve. This deterioration of performance occurred only when the maintenance was transferred from the Navy to the contractor. Thereafter, when

contractors changed, quality did not suffer. This suggests that the problem lay with transitioning from Navy personnel to civilians. After that, when the contract changed hands, only the management team changed. Because the workers remained in place, the benefit of their experience was retained, and the transition was smooth.

In this case, competition/outsourcing worked. The contractors improved the availability of aircraft to the training command while using resources more efficiently. This does not prove that the in-house Navy system was inefficient, but the competition from outsourcing allowed any weaknesses to be exposed and resolved. The downside of this case is that the contractors took a long time to recover from the declines in the maintenance rates during the transition.

# Introduction

During this period of tight budgets and government downsizing, the Department of Defense has looked closely at outsourcing and privatization as a means to reduce infrastructure costs. To achieve the desired levels of ships, aircraft, and troops, and overall modernization, supporting infrastructure must cost less and be more efficient. Outsourcing and privatization are methods of moving functions once done by the government to the private sector, which can save money and resources.

Even if outsourcing makes more sense fiscally, from the warfighters' point of view, service quality, not cost, is paramount. The warfighters view the supporting infrastructure not as a logical place to save money, but as a resource that gives commanders the means to perform their missions, and ultimately, to fight and win wars. Will reducing the infrastructure by outsourcing strengthen or weaken our warfighting ability? Will spending less on infrastructure mean the troops must fight with inferior weapons, training, and equipment or will it give them more resources to do their job?

In this paper, we analyze a particular outsourcing experience to determine whether outsourcing works in a practical sense; that is, does the military really get an equivalent output? This is not a cost comparison; it is a comparison of productivity.

According to the Defense Science Board Task Force on Outsourcing and Privatization, all supporting activities ashore can be outsourced [2]. The services have known for a long time that savings can result from competing and outsourcing functions to the private sector. In the mid 1980s, for example, the Navy turned over the maintenance of aircraft in the training command to private contractors. Here we concentrate on one specific area of that conversion. We attempt to quantify the changes and compare the effectiveness and efficiency of Navy and contracted maintenance.



## Background

The Navy trains its aviators in phases at a number of locations. All future pilots begin primary flight training in a fixed-wing aircraft, the T-34C Turbo Mentor. Upon completing this 22-week program, student pilots are divided into one of three pipelines for further training: strike (jets), rotary wing (helicopters), or maritime (propeller). Those students selected for strike are sent to intermediate training at either NAS Kingsville in Texas, or NAS Meridian in Mississippi for a 23-week course flying the T-2C Buckeye. Those who qualify then take advanced strike training, which is also given at Kingsville and Meridian. On successful completion of this 25-week advanced course, the students earn their wings. Until recently, these students received all their advanced training on the TA-4J Skyhawk.

A-4 Skyhawks are single-seat, single-engine light attack jet aircraft flown by the United States Navy and Marine Corps and a number of foreign military services. They first came into service in the 1950s as nuclear-weapon-capable strike aircraft and were flown extensively during the Vietnam conflict in a conventional attack role. Highly regarded for their exceptional handling qualities, a two-seat version was adapted to serve as the Navy's advanced jet training aircraft. The Navy's Flight Demonstration Squadron (the Blue Angels) flew Skyhawks for many years as well. The last production aircraft was completed in February 1979 and delivered to the Marine Corps. The airframe has been retired from its primary role as an attack aircraft, but some are still flown in the fleet, primarily as adversary aircraft at commands such as the Fighter Weapons School (Top Gun) and composite squadrons, where they fill various roles. The TA-4J version still serves in the advanced training command at NAS Meridian, but it is gradually being replaced by the T-45 Goshawk.

Before the mid 1980s, all the Navy's O- and I- level maintenance on the TA-4J aircraft was done in-house. That is, Navy enlisted personnel performed all required work. Here we refer to this as the "in-house" case.

As a result of an A-76 competition, the Navy outsourced the maintenance of the aircraft. The contractor's winning bid was about 20 percent lower than the in-house bid, after taking contract management and competition costs into account. The work was transferred to the contractor one level at a time, starting in 1985. The overall list is stylized in figure 1. The first contract was awarded to Lockheed for O-level support in 1985, and the contract switched to Grumman in 1990. Burnside OTT won the first contract for I-level support in 1989. In 1993, the contracts for O-level and I-level were combine into one contract awarded to UNC. Finally, D-level support was awarded to Saberliner in 1995.

Figure 1. TA-4J contract maintenance award dates

Year	O-Level	I-Level	D-Level
1980	ORGANIC		
1985			
1986	Lockheed	Burnside OTT	
1987			
1988			
1989	Grumman		
1990			
1991			
1992	UNC		
1993			
1994			
1995			Saberliner

A previous CNA study [3] found that the material readiness of surface ships after depot maintenance was about the same whether the work was done in a public (Navy) yard or a private yard. Our work complements that research. Here we examine O- and I-level maintenance for aviation. O-level maintenance can be completed by the squadron (or its contractor equivalent) and consists of such tasks as troubleshooting discrepancies, basic engine and transmission repair, and replacing parts. I-level

maintenance, the next step up, is completed by the Aircraft Intermediate Maintenance Department (AIMD).

We chose the A-4 for this comparison because we have ample data for much of the time period that the maintenance was done by Navy personnel, and from the time contractors took over maintenance until the present day.

We collected all our data from the CNA Aviation Information Digest (AID) database, which is a compilation of the NAMS0 4790.A7936, Aviation 3-M Data Report. We took the training data directly from training squadron reports before maintenance was shifted to the contractor and from training wing reports after maintenance was outsourced.

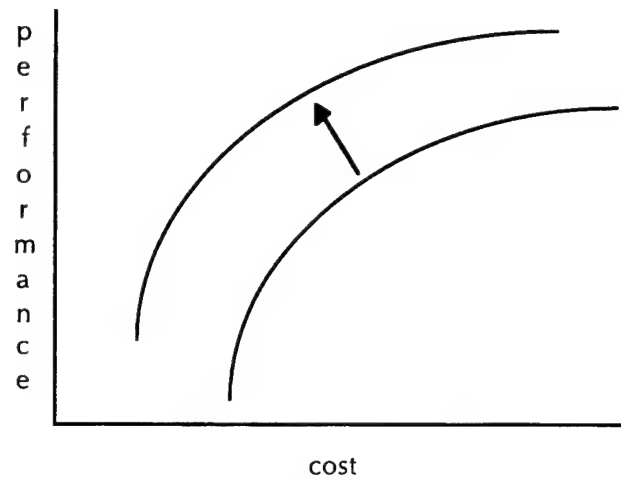
The 3-M data are maintenance statistics reported monthly by all aviation commands to the Navy Aviation Maintenance Support Office. All data are compiled from Maintenance Action Forms (MAFs), which are comprehensive accounts of all maintenance completed on each aircraft in a command. Examples of the statistics include percent full mission capable (FMC), cannibalizations per 100 flight hours, and total number of AIMD parts processed. These data are a known value throughout the fleet, are common across communities, and are easily accessible. To address quality, we examined the following O-level data for the period from April 1980 until January 1996:

- FMC: Percent of time the aircraft is fully ready, with zero system degradation. Higher rates for this category are more desirable, meaning the aircraft are fully ready a larger percentage of the time.
- Mission capable (MC): Percent of time the aircraft can fly and complete a mission. It is more inclusive than FMC, and, as with FMC, higher values are more desirable.
- Direct maintenance man-hours per flight hour (DMMH): The amount of O-level maintenance completed for every flight hour flown. Lower rates can show greater efficiency.

The expected effect of competition/outsourcing is an increase in efficiency (also termed productivity). This is either an increase in the

performance available at any particular cost or a decrease in the cost of attaining a specified performance. Figure 2 shows this expected effect.

Figure 2. Competition/outsourcing increases efficiency

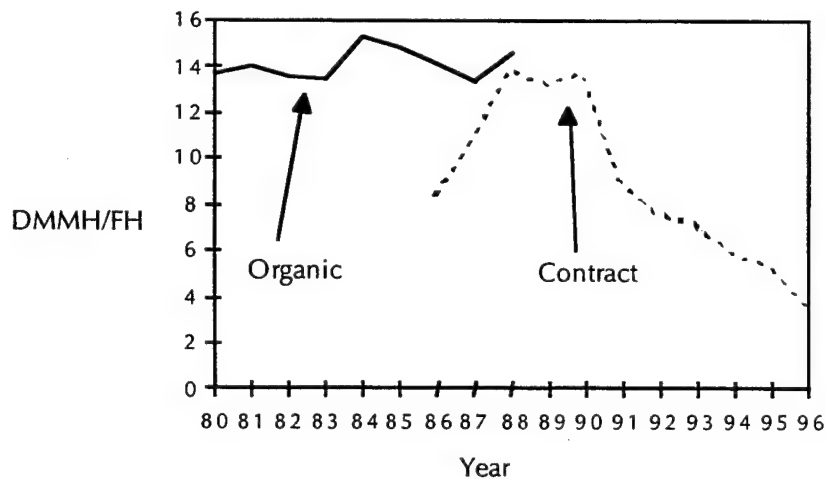


Whether this increased efficiency will result in lower costs, higher performance, or some combination of the two, depends in part on how the contract is written.

## Analysis

The first data set we looked at was direct maintenance man-hours (DMMH) per flight hour, as shown in figure 3. We use the DMMH rate as a proxy for costs. The initial DMMH rates for the contractor rose sharply as the contract went into effect. This rise may indicate that the contractor had to negotiate a "learning curve" or "break-in" period at the start of the contract. Yet even with this initial adjustment, the change in DMMH rate seems to be significant and an improvement over the in-house Navy maintenance. This would suggest that the contractors could ready the aircraft to fly equivalent hours more efficiently (i.e., with fewer personnel). The question then becomes: If the contractor used fewer man-hours than the in-house maintenance team, was overall readiness affected?

Figure 3. O-level DMMH per flight hour

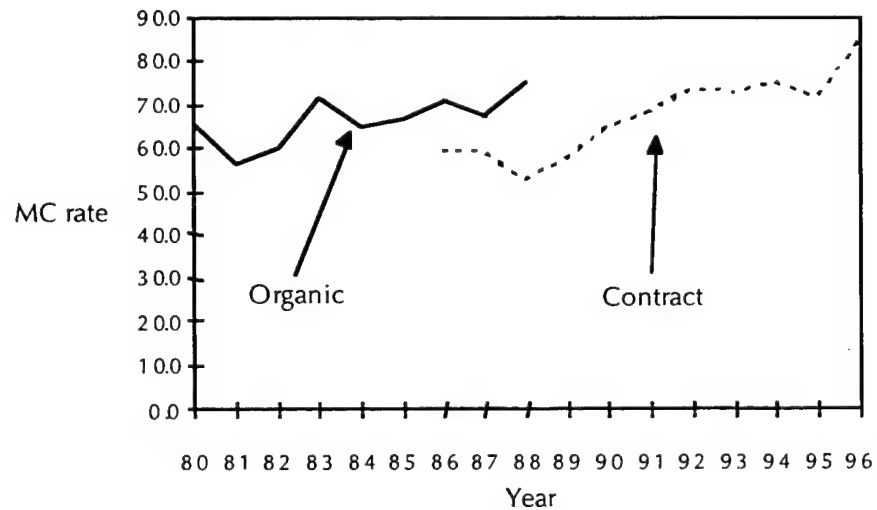


One way to measure the learning curve or break-in period is to determine how long it took the contractor to reach the mean level for that rate. We will see that the mean DMMH rate under contractor maintenance was 9.60, and that it took the contractor 55 months to reach this level. Even so, the contractor's first efforts were better than the in-house DMMH rate. Also, it took the contractor only about 21 months to arrest the rising trend in contract DMMH and begin to lower it to more efficient rates.

Although we based all our conclusions on month-by-month data, the graphs shown here are smoothed average annual numbers, which eliminated the jumps in the plots one might expect with the large standard deviations encountered.

Turning to mission capable rates, the generally positive trend drops off with the start of contract maintenance (figure 4). The rate eventually climbs to a level similar to when the Navy was doing the maintenance. But is it statistically significant?

Figure 4. MC rates

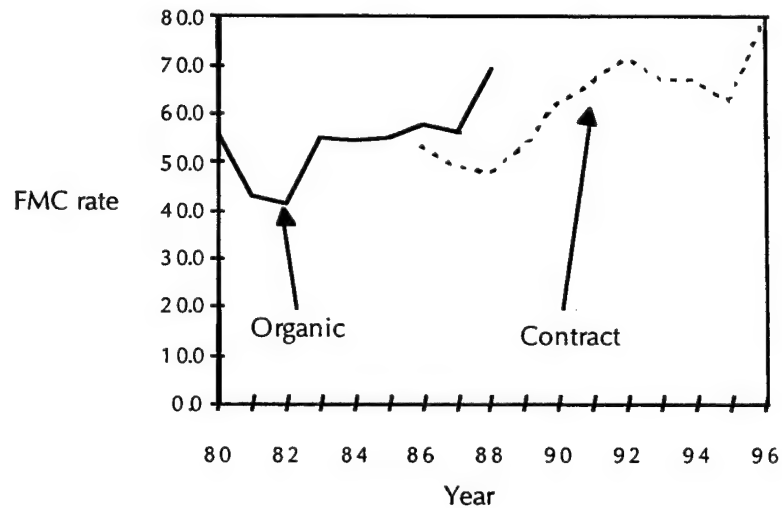


The break-in period for MC rates was nearly as long as for direct maintenance man-hours. The contractors took 49 months to reach the mean MC rate of 65.17 percent. In this case, the contractor also has some ground to make up to reach the level of the in-house team's MC rate. However, the contractor's rate began to improve after about 29 months.

The MC rate is the only 3-M data set that is specified in the actual maintenance contract. Contractor performance which does not meet an MC rate of 65 percent for the TA-4J for 3 consecutive months is reason to end the contract. (The two other rates that the contractor is obliged to meet—aircraft ready for issue at 55 percent and sortie-completion rate at 92 percent—are not included in 3-M statistics.)

The FMC rate graph (figure 5) shows an overall improvement both for the Navy maintenance and the contract maintenance. After recovering from an initial downturn, the rate for contract maintenance eventually surpasses that of in-house Navy maintenance.

Figure 5. FMC rate



It took the contractor 41 months to reach a mean value of 59.81 percent for FMC. About 29 months elapsed before the contractor's rate began to improve.

Table 1 shows the basic information for the formal hypothesis test conducted on the data collected for the in-house Navy and contractor maintenance on the TA-4J Skyhawks of the training commands.

Table 1. Results of z-test

	DMMH		MC		FMC	
	Contract	Organic	Contract	Organic	Contract	Organic
Mean	9.6	14.52	65.17	66.15	59.81	53.47
Std. dev.	3.71	2.37	8.88	6.9	10.05	8.85
Z critical (RR bound)	-1.64		1.64		1.64	
	Lower tail		Upper tail		Upper tail	
Test statistic (z)	-11.68		-0.90		4.89	
Conclusion	Reject $H_0$		Accept $H_0$		Reject $H_0$	

The test rejects the null hypothesis that the DMMH means are equal in favor of the alternative hypothesis that the contract DMMH mean is less than the in-house DMMH mean.

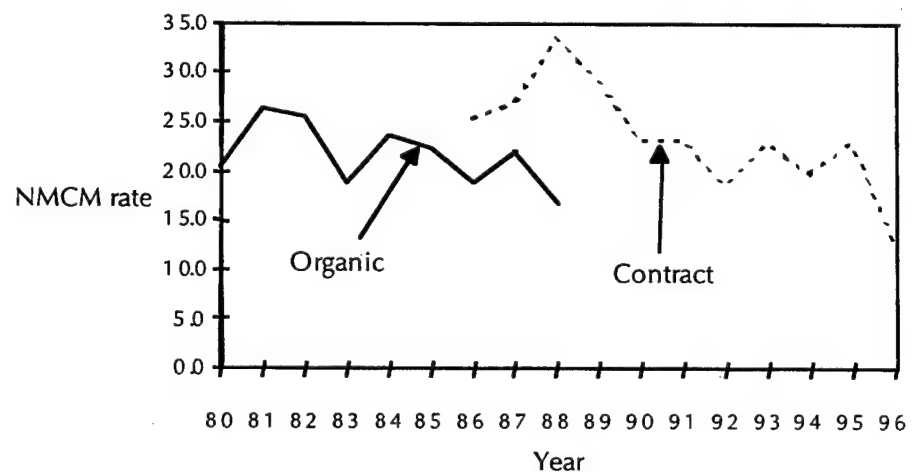
The test accepts the null hypothesis that the MC means are equal for both the in-house and contract cases. The test rejects the null hypothesis concerning the FMC rates. Instead, it accepts the alternative hypothesis that the contract FMC mean is greater than the in-house FMC mean.

This suggests that the contractors had a better FMC rate than the Navy. However, for MC rates, the Navy and the contractor performed about the same. The DMMH rate is interesting. The conclusion is that the contractors had a better DMMH rate than the in-house Navy team. In other words, the contractors used fewer maintenance man-hours to complete a flight hour than the in-house Navy team did. This means that the contractors had the training aircraft ready to fly at rates that equaled or exceeded the Navy's, and the contractor did it with fewer maintenance man-hours.



Although we believe that DMMH, FMC, and MC capture the substance of the comparison, we analyzed other 3M data categories as well. Another data set that can be used to estimate quality of maintenance is the “not mission capable” (NMC) rate. In one sense, this data set is the opposite of mission capable. (Either the aircraft is mission capable or it is not mission capable.) However, NMC status is subdivided by cause: not mission capable due to supply reasons (NMCS) or maintenance reasons (NMCM). An example of NMCS would be a part is needed but is not available, whereas an aircraft would be NMCM if the parts are available but the person with the correct skills is not. To complicate things further, an aircraft can accumulate NMCM time and NMCS time concurrently. Figure 6 shows the NMC data for the periods in question.

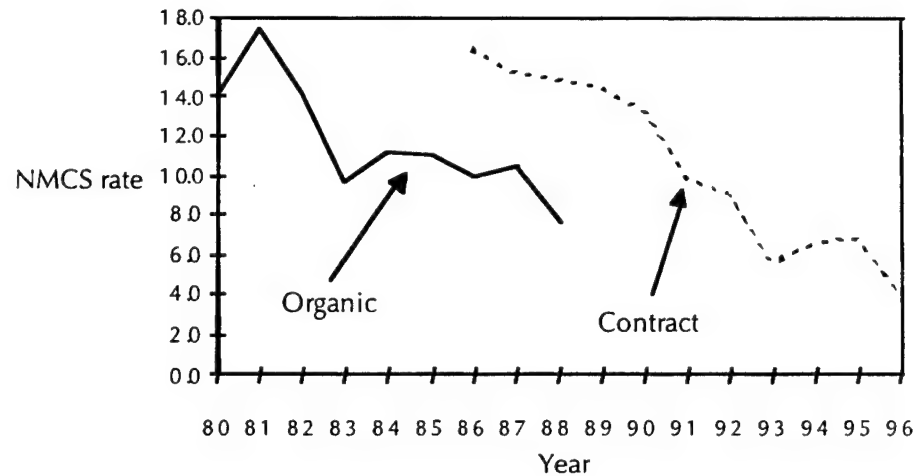
Figure 6. NMCM rate



Notice that the contract NMCM rate immediately jumped to a much higher level than the in-house rate, and then continued to climb. This is an excellent example of the learning-curve effect. Figure 7 shows the results from the NMCS data.

Although the contract rate for NMCS again jumps from the in-house level, the contractor makes a steady and dramatic improvement, even

Figure 7. NMCS rate



when compared with the low rates of the last years of in-house maintenance. In fact, the improvement in the NMCS rate allows the contractor to match the overall NMC rate of the in-house years. Table 2 summarizes the NMC data.

Table 2. Highlighting the NMC rate

	NMCM		NMCS		NMC total	
	Organic	Contract	Organic	Contract	Organic	Contract
Mean	22.31	23.95	12.26	10.87	34.57	34.83
Std. dev.	3.76	5.66	3.26	4.59	6.40	8.88
Z critical (RR bound)	-1.64		-1.64		-1.64	
	Lower tail		Lower tail		Lower tail	
Test statistic (z)	2.49		-2.52		0.24	
Conclusion	Cannot reject $H_0$		Reject $H_0$		Accept $H_0$	

The z-test for the NMC statistics gives some interesting results. The test is a lower bound case because the alternative hypothesis states that the contract means are less than the in-house means. In other words, to support the alternative that the contract means are "better" than the in-house means, we want the test statistic to be sufficiently

negative to be below the critical value, and therefore reject the null hypothesis. If the opposite were true, and the in-house means were greater than the contract means, then the test statistic could be positive (or at least, not sufficiently negative) and support the null. This is true in the case of NMCM. The test statistic is 2.49 with the critical value at 1.64, and normally we could accept the null hypothesis that the contract and organic means are equal. Intuitively, however, it is hard to make that claim. The contract mean is "worse" than the in-house mean and has a higher variance. If the alternative hypothesis is altered to a two-tailed test by stating that the contract and in-house means are not equal (as opposed to saying one is greater or less than the other), the critical value changes to  $\pm 1.96$ , and the test statistic falls in the upper portion of the rejection region. For this reason, and to keep the test consistent with the rest of the study, the conclusion for NMCM as stated in table 2 is that we cannot reject the null hypothesis that the contract and the in-house means are equal.

The NMCS rate is much more straightforward. As table 2 clearly shows, the test statistic falls in the rejection region, and we can reject the null hypothesis that the contract and in-house means are equal in favor of the alternative that the contract NMCS mean is less than (better than) the in-house NMCS mean.

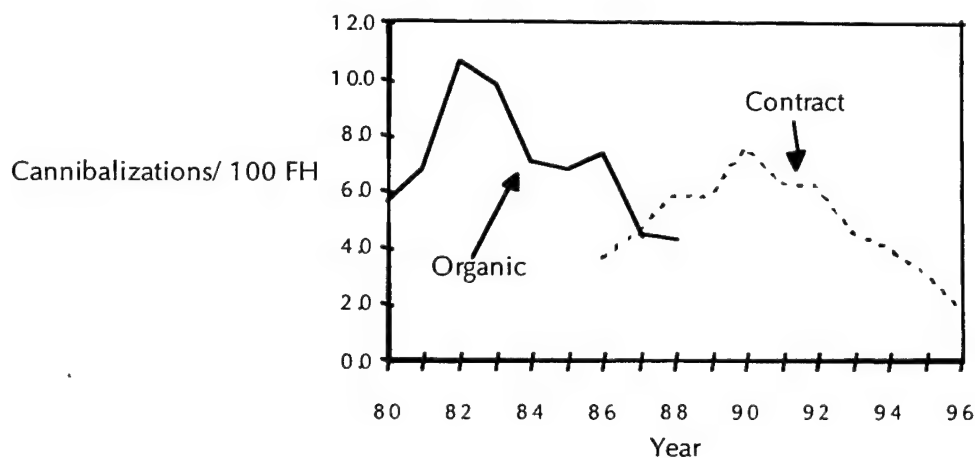
The overall result is that the supply portion of the NMC rate indicates that the contractor rates were better, yet the maintenance portion of the NMC rate seems to show that in-house personnel did a better job than the contractors. Putting the two halves together and recalculating the test implies that the two portions of the overall rate were close enough to wash out any advantage one had over the other, leaving the final conclusion that the in-house NMC rate equaled that of the contractor. This is consistent with the results from table 1. Remember that not mission capable and mission capable are opposites, so if we determined one data set to be equal, we would expect the other to be equal as well.

Why the two components of NMC differed is harder to discern. Why would the contractor have a higher NMC rate due to maintenance factors? One possibility is that the contractor did not initially hire enough workers to do the job in a timely fashion, or did not hire work-

ers who had the needed training. The other possibility is that the contractor did not start out with the right support equipment.

We could argue that to keep mission capable aircraft flying, the contractors cannibalized parts from other NMC aircraft. Cannibalization is the practice of taking supply parts off an aircraft that is “down” (not mission capable) and putting them on another aircraft that may need only one or two parts to keep it flying. It is tracked as the number of cannibalizations per 100 flight hours. A high cannibalization rate may result from inefficiencies in supply or maintenance administration. Figure 8 shows that at first, the contractor did little cannibalizing. Eventually, the contractor cannibalization rate climbed, but then it leveled out and soon declined to a rate that was lower than it had been at the start. The mean rate for the contractor was 5.26, whereas the mean rate for the in-house team was 7.56. Although the contractor’s rate did climb, it never approached the high mark for cannibalization under in-house maintenance.

Figure 8. Cannibalizations



## Intermediate-level maintenance

Most of the previous data has dealt with O-level maintenance. The contract for the TA-4J also covers I-level maintenance, which is work

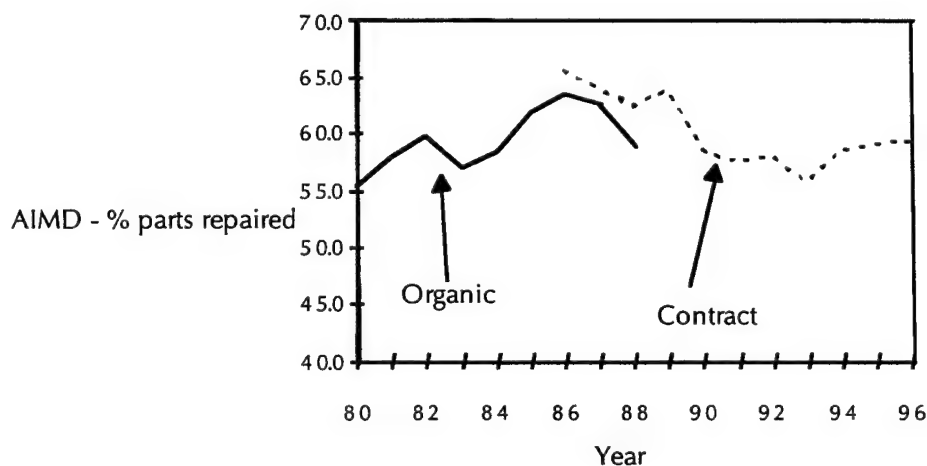
done by an Aviation Intermediate Maintenance Department (AIMD). Three statistics give some insight into this level of work. When a "gripe" (defect) is written against an aircraft or component, one of three things happen:

- The gripe is repaired.
- The gripe cannot be fixed at that level of maintenance. (It is beyond the capability of maintenance (BCM).)
- No defect is found (or the problem can't be attributed to the part being worked on).

Figure 9 shows the percent of AIMD items processed that were repaired.

At first glance, the contractor and the in-house rates appear to be heading in opposite directions. We used a linear regression to conduct a time trend analysis with inconclusive results. Even when we included an apparently significant break-in period, the contractor mean for AIMD parts repaired was 59.9 versus a rate of 59.4 for the in-house team. These are statistically equal.

Figure 9. Percentage of parts repaired by AIMD



The second statistic gives the percentage of AIMD items processed that were declared BCM, that is, items that had to be passed on for depot-level repair. Figure 10 shows this information. Again, the contractor appears to have higher rates compared to in-house maintenance. With a mean of 31.6, in-house maintenance clearly has a lower rate than the contractor mean of 34.8, and in this case, the z-test bears this out. For the scope of this study, the BCM rate might be the least meaningful. We assume that many items processed as BCM is a bad thing, indicating an inability to do the work on a lower level. This may not be the situation at all. By the time the contractors took over the work, the aircraft fleet had aged considerably, and may well have had many defects requiring depot repair. In the last years of in-house maintenance, there was also an upturn in the mean AIMD BCM rate, which may support this idea.

The third statistic is the percentage of AIMD items processed with no defect. (See figure 11.) Finding that the part processed wasn't the cause of the original gripe, or that the part processed wasn't defective are two reasons why a gripe may be in this category. Although both in-house and contract rates were low, the contractor rate showed vast improvement throughout the time period. As the hypothesis test bears out, the contract rate for AIMD no defects is statistically lower than the in-house rate.

These data, although interesting, do not make a simple comparison for quality purposes. I-level maintenance data do not easily relate directly to the problem of keeping aircraft flying. Nor does the contract parameter dealing with I-level relate well to the 3-M statistics we are using. (The contract calls for a 90-percent ready-for-issue rate for aircraft components inducted for I-level maintenance for which C1 or C3 capability exists as determined by the Activity Individual Component Repairability List (ICRL).) Because no data set we looked at in relation to I-level favored either the contract or the in-house style of maintenance, the best we can say is that not much distinguishes the two, and ultimately, contract I-level maintenance is at least as good as in-house maintenance.

Figure 10. Percentage of AIMD items processed that were BCM

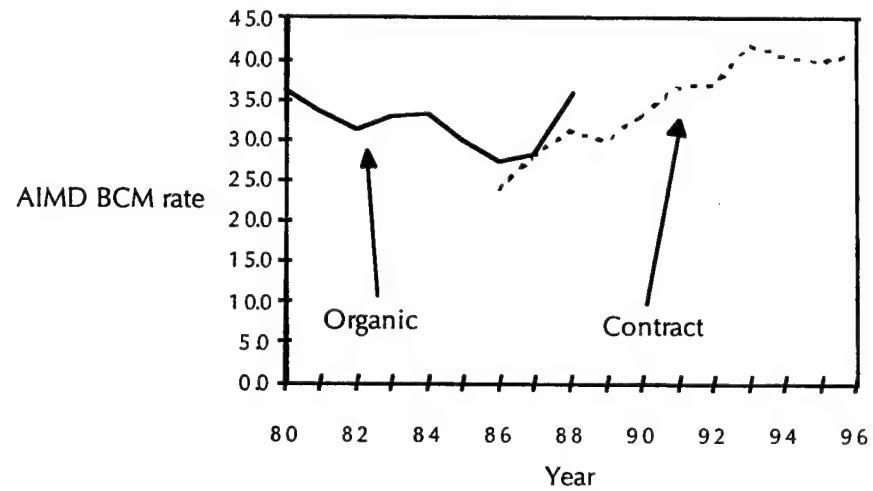
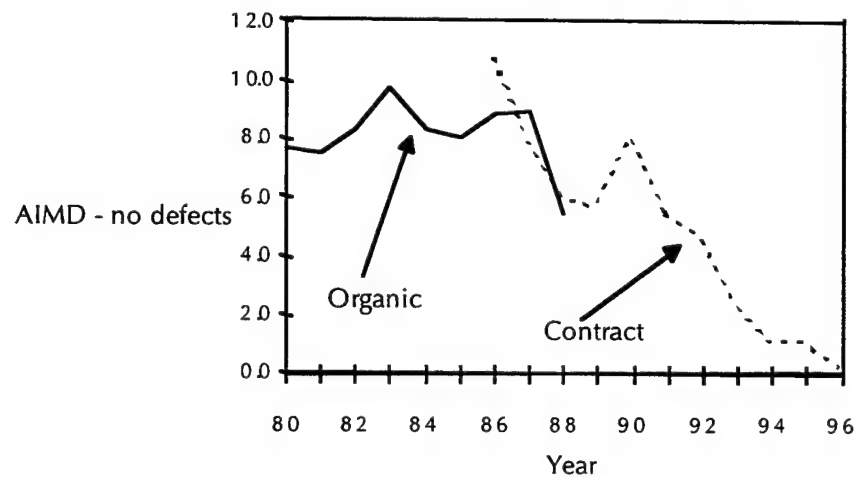


Figure 11. Percentage of AIMD items processed with no defect



## Break-in period

We assumed a break-in period, or learning curve, would be necessary when the contractor first took over. In most of the data sets we analyzed, there indeed seemed to be a period at the start of contract

maintenance when maintenance rates trended down from levels achieved under in-house maintenance. For example, the number of contractor maintenance man-hours started low, but then climbed rapidly to a peak before falling to a low value once again. Explanations for this phenomenon could include inexperienced personnel, insufficient staff, incorrect or insufficient support equipment, or an increase in the amount of work required. Once they were past this initial learning curve, the contractors consistently improved maintenance rates, even when the contract was recompeted and changed hands. Anecdotal evidence suggests that this might have occurred because the same workers remained when the contractors changed, but few government employees joined the initial contractor. In this case, those government employees were mostly enlisted Navy personnel who were ordered to sea-going billets.

To determine how the break-in period affected the comparison between in-house and contract maintenance, we removed those particular months from the contractor's dates and did the comparison once more. Table 3 shows the results from this round of testing.

Table 3. Results of z-testing with break-in periods removed

	DMMH		MC		FMC	
	Contract	Organic	Contract	Organic	Contract	Organic
Mean	9.5	14.5	68.2	66.2	63.4	53.5
Std. dev.	3.96	2.37	7.57	6.9	8.12	8.85
Z critical (RR bound)	-1.64		1.64		1.64	
	Lower tail		Upper tail		Upper tail	
Test statistic (z)	-10.63		1.90		7.92	
Conclusion	Reject $H_0$		Reject $H_0$		Reject $H_0$	

When we remove the break-in periods, we reduced the number of observations from 115 to 94 for the DMMH rate and to 86 for both the MC and the FMC rates. The in-house observations were left at 98. These reductions reflect the months until the rates began to improve.



As we would expect, the contract numbers improve for all three rates. The standard deviations for the contract FMC and MC rates were reduced as well.

In the end, this leads to no real change for the DMMH and FMC conclusions. The mean rates for the contractor are "better" than the corresponding mean rates for the in-house team.

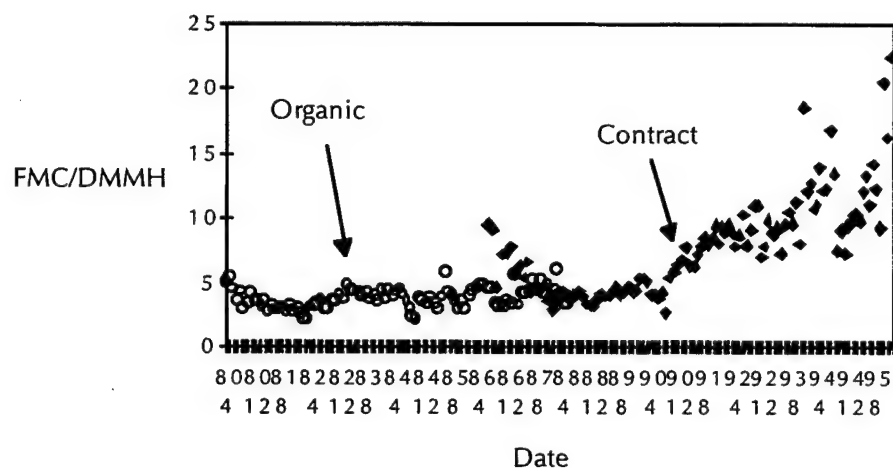
In the original analysis, the conclusion was that mean MC rates for the contractor and the in-house team were equivalent. Eliminating the break-in period does improve the contractor's mean MC rate which rises to 68.2 (from 65.2).

However, statistical tests do not support a conclusion that contractor rates are better than MC rates attained by the in-house maintenance team. The conclusion can be drawn that the contractor's mean MC rate is at least as good as the in-house team's.

## Productivity

Figure 12 shows the ratio between full mission capable and direct maintenance man-hour rates. It shows how much maintenance was needed for each percentage point of full mission capable status. If, for example, we can consider the FMC rate as an output and the DMMH as an input, the ratio is similar to what economists call labor productivity.

Figure 12. Estimate of labor productivity

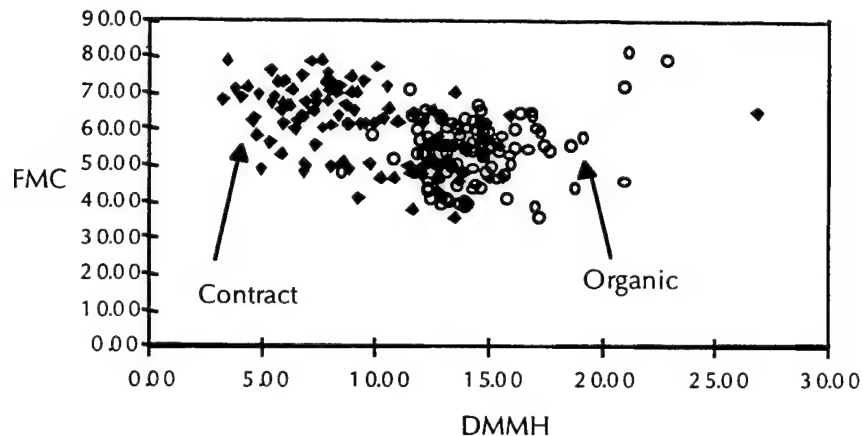


During the in-house period of maintenance from early 1980 to mid 1988, the ratio was fairly steady at around 4 or 5 to one, or, one maintenance man-hour led to 4 to 5 percentage points of FMC status. When maintenance was first contracted out, the ratio increased briefly to about 10, then fell to the in-house level. After a period of "learning," the ratio started to rise, eventually reaching about 15 by 1995.

The evidence suggests that the contractors experienced a break-in period at the start of the contract. After about 2 years, they reversed the downward trend and began to improve. Just how the contractors managed this reversal is still unknown, but possible explanations include capital improvements to the physical plant, increased training, and increased hiring. Again, it is striking that no break-in period is evident through the remaining years of the contracts, denoting a smooth turnover between contractors. Probably, this is because contract workers remained when a new contractor took over. In the end, the contractor was getting at least twice the FMC rate for every man-hour of maintenance completed versus the organic maintenance. Contract maintenance seems to be more efficient.

Figure 13 tells much the same story as figure 12, but in absolute numbers.

Figure 13. FMC versus DMMH



Again we see that when the contractors assumed the maintenance of the TA-4Js in the training command, a higher FMC rate eventually emerged, and the maintenance man-hour rate declined. For now, we can only guess at the reasons. It may be that because the workers often remain in place and do not rotate as Navy technicians do, they gain more experience and efficiency. Or it may be that the contractors

greatly improved the capital of the facilities. Whatever the reason, the contractor seems to be more efficient, and is putting out a "more ready" product at a lower man-hour cost.

One obvious outlier in the contract data corresponds to September 1990 and is characterized by a very high DMMH rate. The FMC rate for that month is unremarkable; however, both the sortie rate (the number of flights) and the utilization rate (the number of flight hours per airframe) for the month are unusually low. September marks the end of the fiscal year, and anecdotal evidence suggests that the training command may have run out of money to fly that month. Yet no other September in the data set jumps out like September 1990. (September 1990 was also the second month of Desert Shield. But neither August nor October 1990 show any unusual variations from other years.)

## **Labor productivity versus technological advances**

Throughout this study, we have contended that we can measure quality and cost of maintenance from the 3-M statistics and that one method (contract or in-house) could be rated as "better." We must understand, however, what this classification does and does not mean.

Because the Navy had a higher DMMH rate than the contractor does not necessarily mean Navy maintenance personnel were inefficient. For the capital outlay the Navy put into the facilities, the in-house maintenance team may have been working at peak efficiency. With no regard for cost, the Navy may have been content to operate with outdated equipment and facilities, as long as it was meeting the minimum requirements.

Private firms, on the other hand will not operate a plant than cannot cover costs. They have an incentive to ensure the plant is running at peak efficiency, even if that means paying for the most advanced technology. The goal is to have a plant that provides the greatest surplus over operating costs. Loosely defined, this is "best practice." Private firms will gravitate to a best practice plant much faster than the Navy could or would. In other words, it is possible that the contractors

increased productivity as a result of technological improvements in the facilities and not because in-house maintenance was inefficient.

From an economist's standpoint, this may well be true. The view of the customer, however, might be much less specific. Outsourcing and privatization save money because they create competition and cost visibility. The result is that the contractor has an incentive to lower input costs and the customer receives better productivity, either through increased output or decreased costs.

## Conclusions

When the transition from in-house to contractor maintenance began, most 3M data rates got worse. The contractors took almost 4 years to reach the previous mean levels and 2 years before they showed any appreciable improvement. Thus, for 2 to 4 years, the training command suffered reduced MC and FMC rates.

In the long run, outsourcing the maintenance function in the training command has not hurt the quality of maintenance, nor has it affected readiness. Because of unanswered questions regarding labor productivity, we cannot say now that the contractors are "better" than in-house maintenance, but we can say that once the break-in period is over, the level of quality provided by the contractors is at least equal to that of the previous in-house team.

Finally, the contractors used far fewer resources, as measured in maintenance man-hours, to complete the job. The contractor provided an equivalent flight hour with a 33-percent reduction in direct maintenance man-hours, an obvious resource and cost savings. Some of these gains could have been used to increase MC and FMC. Outsourcing leads to a gain in efficiency. How this gain is divided between performance gains and cost reductions depends on how the contract is written.

## Appendix: Methodology

We want to compare the 3-M data for both in-house and contract maintenance completed in the training command. By organizing the data and making two simple calculations (mean and standard deviation), we see that the data sets appear to be different. We want to know whether, over hundreds of observations, the data sets are “close enough” to say they are the same, or if they are different enough to say one data set is better than the other. To accomplish this goal, we used a standard z-test.

The objective of a statistical test is to challenge a theorized value for a population parameter. To understand how it works, we must first define some terms:

- Alternative hypothesis ( $H_a$ ): the research theory we wish to support, normally by showing that the converse to the alternative, the null hypothesis, is false.
- Null hypothesis ( $H_o$ ): a statistical theory to be tested and then accepted or rejected in favor of the alternative.
- Test statistic: the formulation used to test the null hypothesis. This is what does the work of the test. For this study, it will be the z-test.
- Rejection region (RR): specifies the values for which the null hypothesis will be rejected or accepted.

As an example of the process, consider the following hypothetical situation. A poll is taken among a random sample of 1,000 Navy enlisted personnel, surveying their satisfaction with bell-bottom dungarees. Forty-seven percent of those polled say they favor wearing bell-bottoms. Can we say that the majority of Navy enlisted personnel dislike bell-bottoms? In this case, taking 50 percent as the point where we reach a majority, we can say our null hypothesis is that, at a minimum, 50 percent of enlisted Navy personnel favor bell-bottoms, that is,

$H_0: p \geq 0.5$ . If we disagree with this assumption, our alternative hypothesis could be that the percentage of enlisted personnel Navy-wide who prefer bell-bottoms is actually less than 50 percent, or,  $H_a: p < 0.5$ . Because the sample size is large, we used the z-test for the test statistic, although no formal proof is offered here. We determined the rejection region by choosing how much error we were willing to accept in our calculations. Here again we need a few definitions:

- Type I error: Rejecting the null hypothesis when it is actually true. The probability of a type I error occurring is represented by  $\alpha$ .
- Type II error: Accepting the null hypothesis when the alternative hypothesis is actually true. The probability of this occurring is represented by  $\beta$ .

Note that  $\alpha$  and  $\beta$  are inversely related; that is, as  $\alpha$  increases,  $\beta$  decreases, and vice versa. The key, then, to balancing the proportion is to ensure that the sample size is sufficiently large to best explain the population as a whole. In general terms, as the sample size increases,  $\beta$  decreases for a fixed  $\alpha$ .

Once the sample size is determined, the most common method of determining the rejection region is to fix the probability of a type I error and look up the boundaries in a table. With the rejection region determined, we can compute the test statistic, and declare one of three outcomes. If the test statistic falls within the rejection region, we can say we reject the null hypothesis in favor of the alternative. If the test statistic does not fall in the rejection region and the value of  $\beta$  is sufficiently low, we can say we accept the null hypothesis. However, to determine  $\beta$ , we must state a specific value for the alternative, and this can be cumbersome. Even if we suggest a valid alternative, calculating the probability of a type II error can be difficult. Therefore, if the test statistic does not fall within the rejection region and we cannot determine the value of  $\beta$ , or, if we can and the value is large, we then say that we fail to reject the null hypothesis, and can try to get additional information before declaring a conclusion.

For the case of A-4J Skyhawk maintenance in the training command, we wish to test whether the means of each contract maintenance data



set are equal to the means of each corresponding in-house maintenance data set. If both data sets are equal, then subtracting one from the other should yield zero. This will be the null hypothesis. Alternatively, we wish to test whether the means of the contract maintenance data set are better than those of the organic data set. Because we have large sample populations, we will use the z-test as the test statistic. The rejection region will be defined using 0.05 as our value for  $\alpha$ , the probability of making a type I error.

$$H_0: (\mu_c - \mu_o) = 0$$

$$H_a: (\mu_c - \mu_o) > 0 \text{ for MC and FMC, or} \\ (\mu_c - \mu_o) < 0 \text{ for DMMH}$$

$$\text{Test statistic: } Z = \frac{(\bar{Y}_c - \bar{Y}_o) - D_0}{\sqrt{\frac{\sigma_c^2}{n_c} + \frac{\sigma_o^2}{n_o}}}$$

$$\text{Rejection region: } z > z_\alpha \text{ (upper tail for MC and FMC), or} \\ z < z_\alpha \text{ (lower tail for DMMH)}$$

Where

$\mu_c$  = the mean values for each contract maintenance data set

$\mu_o$  = the mean values for each organic maintenance data set

$(\bar{Y}_c - \bar{Y}_o)$  = the point estimator of the difference of means

$D_0$  = the hypothesized difference of means, set at 0 in this case

$\sigma_c^2, \sigma_o^2$  = the respective contract and organic population variances  
and

$n_c, n_o$  = the respective contract and organic sample sizes.

We use both the lower tail and upper tail tests because the alternative states that the contract data means are "better" than the organic maintenance data means. For FMC and MC, this means that the contract means are believed to be greater, so the upper tail derivative is suitable. On the other hand, a lower direct-maintenance-per-flight hour rate is believed to be better, so we use the lower tail test.

## References

- [1] Carla E. Tighe et al. *Outsourcing Opportunities for the Navy*, Apr 1996 (CNA Research Memorandum 95-224)
- [2] *Defense Science Board Task Force on Outsourcing and Privatization*, Final Report, Aug 1996
- [3] John D. Keenan et al. *Issues Concerning the Public and Private Provision of Depot Maintenance*, Apr 1994 (CNA Research Memorandum 94-65)
- [4] Robert W. Downey, James M. Jondrow, and Roberta T. Taggart. *Costs of Base Operations in Support of Navy Individual Training*, Oct 1989 (CNA Research Memorandum 89-125)
- [5] Carla E. Tighe et al. *Case Studies in DoD Outsourcing*, Jan 1997 (CNA Annotated Briefing 96-62)
- [6] W. Jevons. *Theory of Political Economy*, 4th edition, 1911

## List of figures

Figure 1. TA-4J contract maintenance award dates . . . . .	8
Figure 2. Competition/outsourcing increases efficiency . . .	10
Figure 3. O-level DMMH per flight hour . . . . .	11
Figure 4. MC rates . . . . .	12
Figure 5. FMC rate . . . . .	13
Figure 6. NMCM rate . . . . .	15
Figure 7. NMCS rate . . . . .	16
Figure 8. Cannibalizations. . . . .	18
Figure 9. Percentage of parts repaired by AIMD . . . . .	19
Figure 10. Percentage of AIMD items processed that were BCM . . . . .	21
Figure 11. Percentage of AIMD items processed with no defect. . . . .	21
Figure 12. Estimate of labor productivity . . . . .	25
Figure 13. FMC versus DMMH . . . . .	26

## List of tables

Table 1.	Results of z-test . . . . .	14
Table 2.	Highlighting the NMC rate . . . . .	16
Table 3.	Results of z-testing with break-in periods removed . . . . .	22

# **Distribution list**

Research Memorandum 97-30

**OPNAV**

N46

N47

N81